The First Peopling of South America: New Evidence from Y-Chromosome Haplogroup Q

Vincenza Battaglia[1], Viola Grugni[1], Ugo Alessandro Perego[1,2], Norman Angerhofer[2], J. Edgar Gomez-Palmieri[2], Scott Ray Woodward[2,3], Alessandro Achilli[4], Natalie Myres[2,3], Antonio Torroni[1], Ornella Semino[1,5]*

1 Dipartimento di Biologia e Biotecnologie "Lazzaro Spallanzani", Università di Pavia, Pavia, Italy, 2 Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah, United States of America, 3 AncestryDNA, Provo, Utah, United States of America, 4 Dipartimento di Chimica, Biologia e Biotecnologie, Università di Perugia, Perugia, Italy, 5 Centro Interdipartimentale "Studi di Genere", Università di Pavia, Pavia, Italy

Abstract

Recent progress in the phylogenetic resolution of the Y-chromosome phylogeny permits the male demographic dynamics and migratory events that occurred in Central and Southern America after the initial human spread into the Americas to be investigated at the regional level. To delve further into this issue, we examined more than 400 Native American Y chromosomes (collected in the region ranging from Mexico to South America) belonging to haplogroup Q – virtually the only branch of the Y phylogeny observed in modern-day Amerindians of Central and South America – together with 27 from Mongolia and Kamchatka. Two main founding lineages, Q1a3a1a-M3 and Q1a3a1-L54(xM3), were detected along with novel sub-clades of younger age and more restricted geographic distributions. The first was also observed in Far East Asia while no Q1a3a1-L54(xM3) Y chromosome was found in Asia except the southern Siberian-specific sub-clade Q1a3a1c-L330. Our data not only confirm a southern Siberian origin of ancestral populations that gave rise to Paleo-Indians and the differentiation of both Native American Q founding lineages in Beringia, but support their concomitant arrival in Mesoamerica, where Mexico acted as recipient for the first wave of migration, followed by a rapid southward migration, along the Pacific coast, into the Andean region. Although Q1a3a1a-M3 and Q1a3a1-L54(xM3) display overlapping general distributions, they show different patterns of evolution in the Mexican plateau and the Andean area, which can be explained by local differentiations due to demographic events triggered by the introduction of agriculture and associated with the flourishing of the Great Empires.

Introduction

Three main events contributed to the genetic structure of modern American populations: the entry into the Americas of Beringian and Asian groups which gave rise to the First Americans, the internal movements determined by the demographic expansions ascribable to the agricultural advances and/or the flourishing of the Great Native American Empires and, finally, the European and African post-Colombian gene-flow. Due to the high rate of admixture between autochthonous people and newcomers from the Old World [1], most of the present genetic information on Native Americans derives from studies of **uniparental markers** [2]. Recently, clues on the origin and spread of the First Americans have been provided by analyses at the level of entire **mitochondrial genomes** [3–12], which revealed a higher level of molecular diversity in the Americas than previously reported.

Several studies have focused on the **Y chromosome** as an independent source of genetic information, but the identification of the male founding lineages has been particularly challenging because of the drastic reduction among modern American populations of autochthonous Y-chromosomes following the male post-Colombian colonization and the high rate (about 17%) of Old World male-mediated admixture [13]. Nevertheless, two founding lineages of Asian ancestry, designated C and Q, were described early on [13–20] and, on average, they account for about 6% and 75% of Native American Y chromosomes, respectively. According to the multimodal distribution of these haplogroups and the different geographic patterns of their sub-clades, their entry into the Americas is likely to have occurred at different times, with haplogroup Q (observed all over the double continent) arriving prior to haplogroup C (limited to North America). During the last two decades, the resolution of these haplogroups has not undergone substantial improvements, with the vast majority of Native American Y chromosomes still falling into one of the two main lineages, Q1a3a1a-M3 [14] and C3b-P39 [13].

As attested by recent mitochondrial DNA (mtDNA) studies, only the dissection of the most common Y-chromosome founding haplogroups, together with major sampling efforts, can permit the

identification of geographically-restricted lineages able to provide information not only on the migration routes but also on the demographic dynamics of the male component associated with the population expansions that occurred at different times and locations in the American continent. With this aim, more than 400 Native American unrelated male subjects, harboring Y chromosomes belonging to haplogroup Q were identified among more than 2,000 subjects of paternal Native American ancestry from a broad geographic area, ranging from Mexico to South America. These subjects have been examined for 23 haplogroup Q-specific SNPs and 37 Y-STR markers. In addition, in the attempt to identify possible genetic sources for Paleo-Indian Y chromosomes, the results obtained were analyzed in comparison with those of different populations, geographically or culturally related to the proposed ancestral Asian source areas, some of them already published, others (the Koryaks from Kamchatka and the Mongols) analyzed in the present study.

The first peopling of America: state of the art

Since 1590, when Friar José de Acosta first argued that the indigenous inhabitants of the Americas likely came from Asian populations [21], the events that characterized this process of colonization of the Americas, the last major landmass to experience human dispersal, have been a matter of debate. Scholars of different fields, including archaeologists, anthropologists, linguists and geneticists seek to answer questions concerning the number of migrations, the homelands of the peoples involved in the process and the routes they followed, as well as the timing of these events.

The traditional model of colonization, first advanced by Hrdlička [22] and subsequently modified in order to incorporate new chronological and cultural data [23,24], assumed a single rapid migration (containing all founding lineages) of Asian ancestors across the Bering land bridge around 16.5 thousand years ago (kya – calendar years BP) [25], the earliest period for which there are significant archaeological findings on both side of Beringia. The basis of this model, called the ''blitzkrieg'' model, was the identification of the Clovis culture, the oldest (~13 kya) and most widely distributed archaeological tradition in North America [26]. Later, Greenberg et al. [27] combining linguistic, archaeological and biological evidence, proposed the ''three-wave'' model according to which the three proposed Native American linguistic groups (Amerind, Na-Dene, and Eskimo-Aleut) entered the Americas subsequently to the Last Glacial Maximum (LGM, ~20 kya), in three chronologically separate waves of migrations from different Asian regions. Afterwards, different models were advanced to explain observations based on new genetic data: some were in agreement with the three waves of migration, [28,29] whereas others proposed a single early entry into the Americas [30,31] or two waves of migration, the first from Central Asia (Altai-Sayan and/or mid-lower-Amur) region prior to LGM, and the second, after the LGM, from Amur-Mongolia-Manchuria [32]. Recently, after extensive consideration of the linguistic classification and the genetic differentiation of the three Native American family groups [2,33,34], the latest genetic screening of the nuclear genome supports a model with at least ''three migration-waves'', with a major contribution from the first arrival, accounting for 100% of the Amerindian ancestry in most Native American populations and with decreasing contributions from the subsequent two migrations, which mainly affected Na-Dene and Eskimo-Aleut groups [35]. In addition, the presence and high frequency of alleles and haplogroups unique to Native Americans (such as the autosomal 9-repeat at microsatellite locus D9S1120 [36,37], the Y-chromosome haplogroup Q1a3a-M3

[20,38,39] and the pan-American mtDNA haplogroups A2, B2, C1b, C1c, C1d, D1, and D4h3a [3–7]) support the scenario that Beringian/Eastern Siberian glacial refuges existed and played a major role in shaping the ancestral gene pool of Native Americans. Here, founder lineages differentiated from their Asian sister-clades by genetic drift prior to entering North America. Then, they followed a swift migration southward throughout the rest of the double continent [3,7,37]. Some have proposed a more complex model, in which the initial differentiation from Asian populations ended with a moderate bottleneck in Beringia during the LGM [5]. Thus, a subsequent remarkable population expansion occurred with rapid settlement along a coastal route, the only one available at that time (from 18 to 15 kya), shortly before the development and spread of the Clovis culture [40]. Recently, through the analysis of the distribution of two rare Native American mtDNA haplogroups, a second major path of migration was identified along the interior ice-free corridor between the Laurentide and Cordilleran ice sheets, more or less concomitantly (17–15 kya) to the principal migration along the Pacific coast [6,41].

Regarding South America, taking into account the archaeological evidence of pre-Clovis cultures, such as Monte Verde in Chile (~15 kya) [42] and the Oregon cave (~14 kya) [43], it is likely that people arrived in the Southern Hemisphere no later than 15 to 14 kya [42,44]. Nevertheless, archaeological evidence (tools and skills for harvesting resources from the sea, such as fish, shellfish, and sea mammals) of sufficient antiquity to account for the first arrivals in southern North America is limited to the Paleo-coastal sites in the California Channel Islands, dated 12–11 kya [45]. Simulation models and archaeological records [42,46] suggest the initial settlement along the coasts was followed by colonization along rivers while recent analyses of mitochondrial genomes from the Southern Cone of South America point strongly to a trans-Andean colonization of the continental interior after the southward coastal migration [9–11].

The first inhabitants of the American continent were hunter-gatherers and fishers, as witnessed by archaeological records [47]. Climatic changes and subsequent environmental changes pushed humans and animals towards new areas with favorable conditions to which, in spite of highly differing environments, they adapted rapidly thanks to the relative homogeneity of available prey along the whole American continent. As already hypothesized for Asia and Europe, there was most probably demographic growth just before the invention of agriculture, which drove humans to develop new technologies to meet the growing need for food. However, this process was relatively slow compared to Eurasia, due to differing and unique environments requiring specific technologies that could not be easily exported into neighboring territories. Ancient cultivations (more than 10–8 kya) [48] were identified only in present-day Mexican territories and in the western side of South America, the locations of the largest American Empires such as the Maya – the longest-lasting civilization in Mesoamerica (~3,500–500 ya), and the Aztecs and Incas (~1,000–500 ya). However, hunting and fishing were maintained elsewhere, giving rise to highly structured societies. The Aztecs, who came to the South of present day Mexico from dry uplands of the North, achieved hegemony when the Mayan Empire collapsed. At the same time, in South America, the Incas were able to extend their control 2,500 miles along the Andes, ruling up to 20 different language areas, until the arrival of the Spanish invaders. Investigating the gene pool of these regions, which cover today at least a portion of Ecuador, Peru, Bolivia, Chile and Argentina, along with Mesoamerica, could provide

important insights into the migration flow and expansion which, from Beringia, eventually arrived to Central and South America.

Results

As reported in Table 1, 463 Y chromosomes belonging to haplogroup Q (M242-positive) were identified in this survey. All of them were analyzed hierarchically for signature markers of all haplogroup Q sub-lineages and for new informative variants identified while typing. The analyzed markers and the relative results are shown in Figure 1, which also illustrates the phylogenetic position of the four new lineages.

New lineages

In the most updated phylogeny of haplogroup Q shown in Figure 1 [49], the four new markers define two new M3 sister lineages (PV3 and PV4) and two M3 sub-lineages (M557 and PV2). M557 was identified while typing by sequencing the M194 marker and consists of a TAC deletion in a TACTAC repeat of its amplicon; PV2, PV3, and PV4 are all T to C transitions discovered while sequencing the M242, L330 and L54 markers, respectively (for details see Table S1). The haplogroup Q Y-chromosome sequenced by Wei and colleagues [50] does not harbor any of these new markers.

Haplogroup	Native Americans N=436	Mongols N=22	Koriaks N=5
Q*			
Q1*			
Q1a*			60.0
Q1a1		13.6	
Q1a2		13.6	
Q1a3*			
Q1a3a*		4.5	
Q1a3a1*	16.7		
Q1a3a1a*	76.8		40.0
Q1a3a1a1			
Q1a3a1a2	0.2		
Q1a3a1a3			
Q1a3a1a4	1.4		
Q1a3a1a5	1.6		
Q1a3a1a6	0.2		
Q1a3a1b	2.1		
Q1a3a1c		68.2	
Q1a3a1d			
Q1a3a1f	0.2		
Q1a3a1g	0.2		
Q1a3b			
Q1a6			
Q1b	0.5		

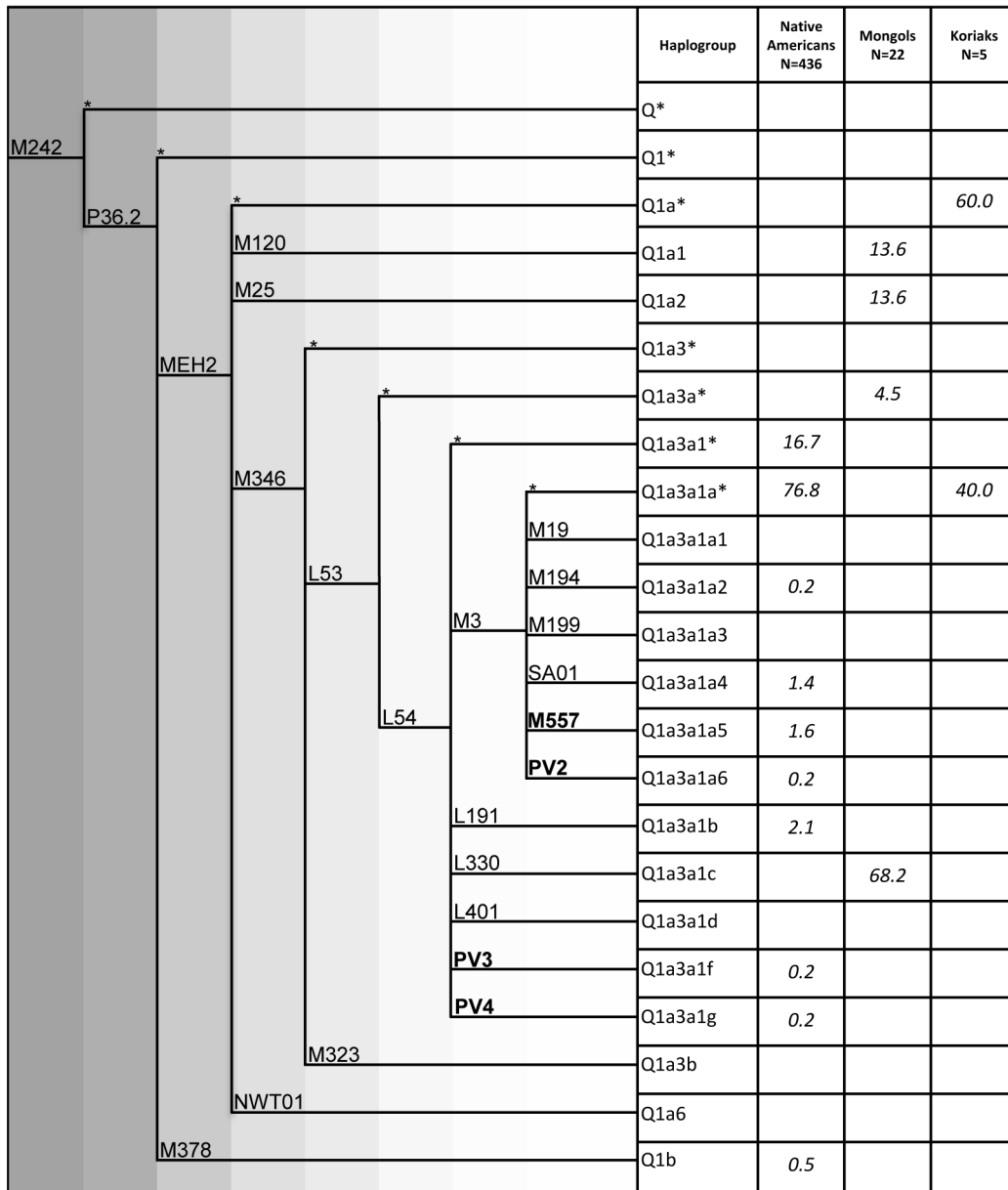Figure 1. Phylogenetic tree of Y-chromosome haplogroup Q. Haplogroup nomenclature and frequencies (%) of haplogroup branches in the analyzed populations (new markers are in bold). The haplogroup labeling is in agreement with the YCC [85] conventions and recent updates [49,51]. *Paragroups: Y chromosomes not defined by any phylogenetic downstream-reported and -examined mutation.
doi:10.1371/journal.pone.0071390.g001

Table 1. Observed Frequencies of Haplogroup Q and of its Sub-Lineages.

Macro-area	Country or Region (N)	Q (N)	Q1a-MEH2	Q1a1-M120	Q1a2-M25	Q1a3a-L53*	Q1a3a1-L54*	Q1a3a1a-M3*	Q1a3a1a2-M194	Q1a3a1a4-SA01	Q1a3a1a5-M557[a]	Q1a3a1a6-PV2[a]	Q1a3a1b-L191	Q1a3a1c-L330	Q1a3a1f-PV3[a]	Q1a3a1g-PV4[a]	Q1b-M378
Asia	Mongolia (1,145)	22		3 (13.6)	3 (13.6)	1 (4.5)								15 (68.2)			
	Kamchatka (34)	5	3 (60.0)					2 (40.0)									
Mexico	Mexico (485)	181					48 (26.5)	123 (68.0)					8 (4.4)		1 (0.6)	1 (0.6)	
Central America	C American Isthmus[b] (14)	3						3 (100.0)									
	Nicaragua (10)	2					2 (100.0)										
	Costa Rica (35)	1						1 (100.0)									
	Panama (242)	37					2 (5.4)	34 (91.9)									1 (2.7)
Andean Regions	NS America[b] (54)	9					2 (22.2)	6 (66.7)									1 (11.1)
	Peru (397)	159					17 (10.7)	126 (79.2)	1 (0.6)	6 (3.8)	7 (4.4)	1 (0.6)	1 (0.6)				
	Bolivia (22)	9					1 (11.1)	8 (88.9)									
	Chile (206)	20						20 (100.0)									
South East America	Brazil (485)	7						7 (100.0)									
	SS America[b] (18)	7					1 (14.3)	6 (85.7)									
	Uruguay (37)	1						1 (100.0)									
Caribbean Region	The Caribbeans[b] (20)	–															
	Total (3,204)	463															

[a]Marker described here for the first time.

[b]Grouping due to sample size <10 in single populations. Central American Isthmus: El Salvador, Guatemala and Honduras; Northern South America: Colombia, Ecuador, Venezuela; Southern South America: Argentina and Paraguay; The Caribbeans: Barbados, Bermuda, Cuba, Dominican Rep., French Guiana, Grenada, Haiti, Jamaica, Martinique.

The following mutations were surveyed in all potential carrier samples but they were not observed: NTW01, M323, L401 and M19. M199 was only typed in a subset of samples.

In parentheses, % frequencies.

doi:10.1371/journal.pone.0071390.t001

Haplogroup distribution

The uneven distribution of haplogroup Q1a3a1-L54 in Native American and Asian samples recently described by [49] was confirmed. This clade harbors virtually all (99.5%) the Native American Y chromosomes. However, when taking into account also the Altaian samples of Dulik et al. [51], it accounts for only 51% of Q Y chromosomes from Asian populations living in the regions that have been proposed as sources and pathways of Native American migrations (Figure 1, Table 1 and Dulik et al. [51]). In addition, with the only exception of two Koryak samples that were M3-positive, none of the observed lineages are shared between Native Americans and Asians, 80% of the first falling into the Q1a3a1a sub-clade characterized by the biallelic marker M3 and 49% of the second harboring the Q1a3a1c sub-clade defined by the L330 marker.

Only two American Y chromosomes did not cluster into the L54 sub-branch. They were both M378-positive, thus belonging to Q1b, a finding never previously reported for Native Americans. Considering that the phylogeography of this infrequent haplogroup is restricted to South West Asia [52–55], the most likely interpretation of this outcome is that they represent an arrival from Asia in contemporary history. For this reason the two Y chromosomes were not included in subsequent analyses.

Focusing on the Native American-specific lineages, Q1a3a1a-M3 is prevalent in all considered populations. Q1a3a1-L54(xM3) has mainly been observed in Central America as paragroup Q1a3a1-L54*, while Q1a3a1b-L191 was observed in eight Mexicans and one Peruvian sample, and Q1a3a1f-PV3 and Q1a3a1g-PV4 were found in single Mexican subjects. Q1a3a1d-L401 was not observed in our sample. Regarding M3 sub-branches, Q1a3a1a4-SA01, recently described in 16 subjects from the Andean region [56], and the new Q1a3a1a5-M557, were observed in 1.5% (N=6) and 1.8% (N=7) of the Peruvian samples, respectively; lastly, Q1a3a1a2-M194 and Q1a3a1a6-PV2 were found in single Peruvian subjects while the previously described Q1a3a1a1-M19 and Q1a3a1a3-M199 (only typed in a subset of samples) were not observed in this survey.

STR haplotype distribution among the Native American haplogroup Q sub-lineages

Figure 2 illustrates the two principal coordinates (PCos) of the 33 STR *loci* variation in the 433 Native American Y-chromosome haplogroup Q samples of the present study (Table S2) in relation to geographic (panel A) and haplogroup (panel B) distributions. The first principal coordinate sharply distinguishes an Andean cluster from the bulk of samples (panel A), while the second coordinate separates the haplotypes related to the M3 background from haplotypes associated with the L54(xM3) context (panel B). Two other well-differentiated groupings stand out from the plots: one harbouring Central American M3 Y chromosomes and the other encompassing a group of Mexican L54(xM3) samples. The location in the plot of the samples carrying the new variants is in agreement with this distribution, with the M557 lineage restricted to the above-mentioned Andean cluster.

High values of haplotype diversity (*h*) were observed for both founding haplogroups L54(xM3)(0.9744) and M3 (0.9915). When geographic areas were considered, L54 displayed high values in Mexico (0.9492) and the Andean (0.9905) (or Peruvian) regions; differently, M3 displayed the highest value in Mexico (0.9897) and a general slightly decreasing north-to-south diversity pattern. In the Far East Asians, M3 showed the lowest value (0.9111) (Table S3 and data not shown).

Network analyses and time estimates

To compare our results with literature data from North America and informative Asian regions, in spite the large number of STR *loci* typed in the present study, network analyses and time estimates of haplogroup Q sub-lineages were performed on the subset of *loci* widely used in evolutionary studies and with well-defined evolutionary mutational rates [57,58]. The median-joining networks of the seven STR haplotypes within the observed Q sub-haplogroups (Table S4) are shown in Figure 3, whereas the estimated ages are reported in Tables 2 and 3. Interestingly, the two founding lineages display an early concomitant arrival in Mexican and Andean regions, when the above-mentioned clusters, ascribable to later expansions, were disregarded.

The network of Q1a3a1a-M3* Y chromosomes includes North, Central and South American Y-chromosomes in a star-like shape, in agreement with an early entrance into America and wide expansion of this lineage in the entire continent. The central haplotype, which includes two Far Eastern Asian samples (one Koryak and one Even [59]), is most represented by Mexican and Andean subjects while Canadian samples are mainly observed as one and two steps derivative haplotypes. Regional expansions are present in all areas but are particular evident in the Andean and Canadian regions.

The network of the Q-L54* Y chromosomes is characterized by multiple reticulations, which suggest that they are part of different sub-clades still to be discovered. In spite the low frequency of these Y chromosomes, their preferential Mesoamerican distribution emerges. None of the L54* Y chromosomes is observed in Far Eastern Asian populations living along the entry route to the Americas; its sub-clade L330 underwent expansion in Mongolia and Southern-Altai, whereas the upstream intermediate L53* characterizes Northern-Altai (Figure 3), with a single observation in Mongolia.

Discussion

Although numerous archaeological, linguistic and genetic studies have been performed to date, many aspects of the peopling of the Americas are still largely unresolved. Due to the extensive non-Native American admixture, which began with the arrival of Europeans after 1492, the main contribution to the reconstruction of Native American genetic origin and history has been provided by uniparental systems, which are not affected by the reshuffling effect of recombination. Unfortunately, the high historical rates of non-native male-mediated admixture into Native American communities have slowed down the investigation from the Y-chromosome perspective, which is essential for a complete picture of population movements and demographic events. This paper, in order to address this deficiency, investigates the genetic history of Central and South America using Y-chromosome variation and focusing exclusively on Y-chromosome haplogroup Q, which characterizes the great majority of modern Native Americans and is virtually the only native haplogroup in South America.

Two main founding lineages, Q1a3a1a-M3 and Q1a3a1-L54(xM3), have been observed in this study. Their variation in the Americas is not as restricted as was hitherto thought, giving genetic support to the Pleistocene/Holocene cultural variability in South America recently proposed by Neves and collaborators [60]. They display overlapping general distributions, but their diffusion patterns along the continent suggest they underwent different demographic processes. The most frequent, Q1a3a1a-M3*, shows a network which encompasses North, Central and South American chromosomes in a star-like shape indicative of a first general north-to-south diffusion followed by local expansions in
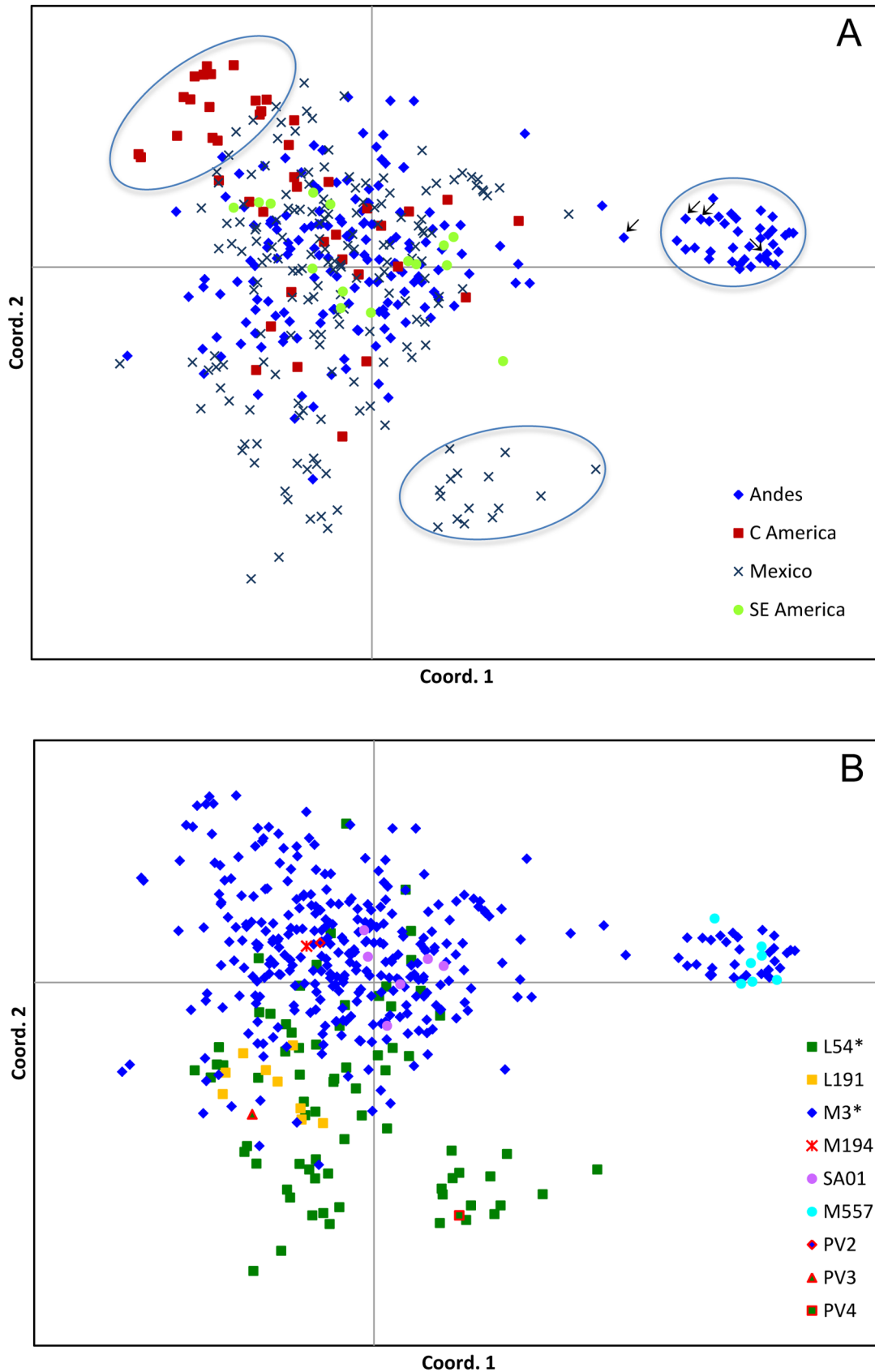
Figure 2. PCoA plots. Analyses were performed on 433 Native American haplogroup Q Y chromosomes and based on pairwise, individual-by-individual genetic distances generated from 33 *loci* Y-STR haplotypes (Table S2) represented in their (A) geographic (<u>Andes</u>: Bolivia, Chile, Colombia,

Ecuador, Peru; Central America: Costa Rica, El Salvador, Guatemala, Nicaragua, Panama; South East America: Argentina, Brazil, Paraguay, Uruguay) and (B) sub-haplogroup Q contexts. Ancient DNA (aDNA) matching haplotypes [75] are indicated by arrows. Three separated (Central American, Mexican and Andean) clusters are indicated by ovals.

Canada, Mexico and the Andean regions in agreement with a general north-to-south decreasing variance (V) and haplotype diversity (h) (Table S3). On the other hand, the network of Q1a3a1a1-L54* is mainly represented by single observations of well differentiated haplotypes and is characterized by many reticulations, indicative of the coexistence of deeper-rooted sub-lineages, with signs of expansions only in Mexico. The presence of these "expanded" haplotypes can explain the lower h encountered in Mexico (0.9344) in comparison with that observed in the Andean region (0.9883) in spite of the higher V (0.518 vs 0.451).

It is worth mentioning that, in these regions, important changes in environmental conditions were brought about by the development of different agricultural/animal domestication practices. In particular, while the subsistence economy of pre-Colombian Mexican society was strictly based on maize cultivation, in the Andes it was more varied, due to the presence of many different habitats and, unlike in Mexico, also took advantage of animal domestication. The consequent creation of ecological niches under different selection pressures may have determined changes in allele/haplotype frequency, as documented by the results of different studies on Central and South American populations [61–63]. In this scenario, even though Y-chromosome haplogroups may not be directly involved in activating this process, their frequency and variation could be influenced by genetic drift in areas where these evolutionary events took place. Accordingly, the different genetic structures of the Mexican and Andean areas, which emerge from the network analysis, are confirmed by the geographic distributions of their sub-clades, which reveal two distinct centers of evolution (Table 1), those rooted in the deeper L54 branch being associated with the Mexican population only, and those rooted in the M3 branch associated exclusively with the Peruvian (Andean) population.

With the exception of only one Q1a3-M346*(xL53) North American sample, all the Native American Q Y chromosomes sub-classified to date (N = 500) for L54 belong to this branch ([49] present paper). Therefore, Q-L54(xM3) should represent (at least in the southern part of the American continent) virtually all the samples previously classified as Q1a3-M346*/Q-M242*, and thus, be classified as the second most important Y-chromosome pan-American founder lineage. The frequencies observed in our population samples are in agreement with a distribution mainly along the northwest border of South America [20,64,65], as proposed in Bailliet et al. [66]. Of the L54 sub-clades, Q1a3a1b-L191 seems generally confined to the Western American coast. Our results reveal its presence in Mexico (1.6%) and Peru (0.3%) and, interestingly, its expansion time (3.6±1.6 kya) overlaps in Mexico (2.3±1.0 kya), coinciding with the historical timeframe of the Mayan Empire. Q1a3a1f-PV3 and Q1a3a1g-PV4 were identified in single Mexican subjects. However, Q1a3a1f-PV3 was found in association with a micro-variant (13.2 partial repeat) at the DYS19 locus, which on the YHRD database (http://www.yhrd.org – release 41 built on December 20, 2012 performed on 105,494 haplotypes within 782 populations) characterizes five subjects, three of them found in the Guatemalan-Mayan population [67], and two in Hispanic American admixed populations from United States [68]. The estimated age for chromosomes carrying that "partial repeat" allele dates back to the Pleistocene/Holocene boundary (Table 2) and, even though characterized by a large standard error, indicates an additional

Mesoamerican differentiation of the most diffused L54. A further potential L54 branch could be represented by the peculiar DYS391–6 repeat allele, which characterizes two subjects from Panama and Colombia of the present study and seven Q(xM3) samples from Central America, five from Panama [69] and two from Nicaragua [70]. The absence in the large dataset of South American Y chromosomes carrying this microsatellite variant would suggest that this lineage was not involved in southern migrations. However, although in the past rare/partial repeat alleles within Y-STR loci were useful for assessing regional Y-chromosome variation or historical migrations and/or refining the resolution of the Y-chromosome tree [71–74], the phylogeography of these identified lineages must be confirmed by surveying a larger set of Native American samples.

The most diffused Q1a3a1a-M3 lineage appears to be more highly differentiated in the Andean region, where new sub-lineages were observed (Q1a3a1a4-SA01, Q1a3a1a5-M557, Q1a3a1a6-PV2), two at polymorphic frequencies. The SA01 marker was observed in Peru and Bolivia ([56], present paper) with an estimated variation age (7.2±2.6 kya, Table 3) compatible with the earliest appearance of agriculture in the Andean region. On the other hand, demographic events promoted by the diffusion of agriculture could explain the local differentiation and expansion of the Q-M3 Andean haplotype cluster, revealed by PCoA (Figure 2), whose estimated age has been evaluated at 4.0±1.2 kya (data not shown) and which includes the younger Q1a3a1a5-M557 sub-clade (2.6±1.4 kya, Table 3). In this context, it is worthwhile noting that the Y-chromosome haplotypes observed in eleven pre-Colombian male individuals buried in four different graves of the Peruvian archaeological site Tompullo 2 [75] overlap the modern haplotypes of the above mentioned M3 Andean cluster (Figure 2), thus providing evidence of a genetic continuity between extant Peruvian and pre-Colombian inhabitants. Taking into account that the Tompullo 2 site likely represents a short-lived settlement of Inca origins that existed during the Late Horizon (15th–16th century) within the context of the mitmaq system (the colonization procedure adopted by the Incas to exploit new resources in new territories, based on movements of small population groups), the above mentioned Andean cluster, which is well separated from the bulk of the Q-M3 Andean Y chromosomes, could be the result of founder effects followed by an additional local expansion in a favorable ecological niche. Worth of note is the absence in our large Andean sample of the lineage Q1a3a1a1-M19 previously described in Native American groups from Colombia [20,76] and recently observed in the Toba group from Argentina [77]. This observation indicates that the continental interior colonization of South America did not occur exclusively through the Pacific coastal route, but was followed by a trans-Andean crossing.

Regarding the Asian distribution of the two Native American Q founding lineages, Q-M3* has been found in different Far East Asian groups, whereas, at present, only a potential Q-L54* has been observed in one Chukchi from Northern Siberia [78]. However, since this L54 sample has not been classified for the L330 marker, which characterizes the Asian branch of Q-L54, and occupies a position close to the Mexican haplotypes (not shown) when included in the L54* network, at present any interpretation of this result (new Asian lineage, remnant of an ancestral state, trace of forward or back-migration) is premature. On the contrary, some information can be obtained by the analysis of Asian M3 Y
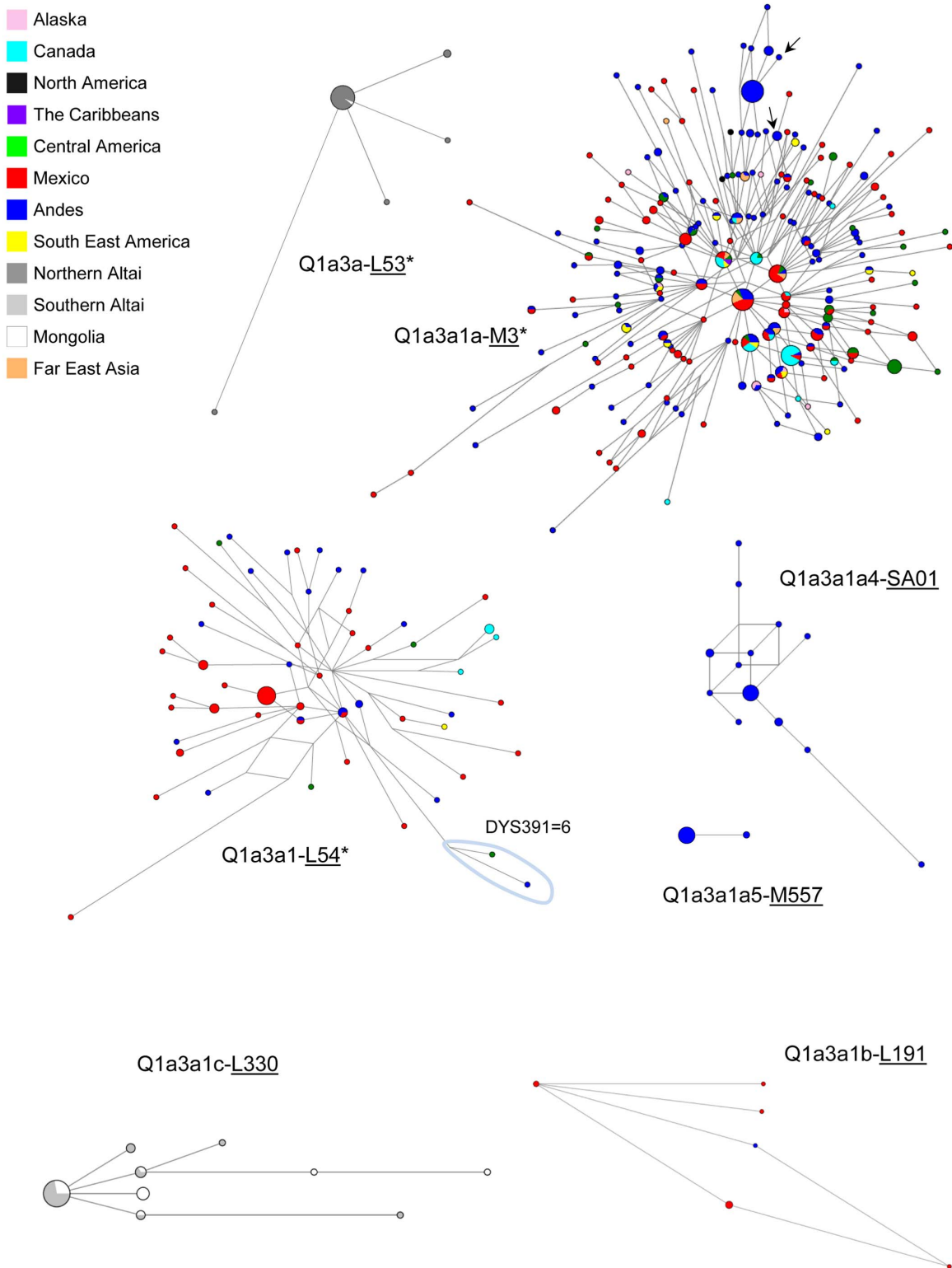
Figure 3. Network of 7 *loci* Y-STR haplotypes belonging to Q sub-haplogroups. Network analysis were performed on Native American [Alaskan, Canadian, North American, Caribbean [49], Central American (from Costa Rica, El Salvador, Guatemala, Nicaragua, Panama), Mexican, Andean (from Bolivia, Chile, Colombia, Ecuador, Peru), South Eastern American (from Argentina, Brazil, Paraguay, Uruguay) (present study)] and Asian [Northern and Southern Altaians [51], Mongolian and Far Eastern Asian (from Kamchatka, Chukotka and Sea of Okhotsk coast) [59,78]; present study)] samples (Table S4). Light blue loop is referred to the unusual DYS391-6 repeat allele. Arrows indicate the 7 *loci* haplotypes matching aDNA [75]. Within the single network, circles and colored sectors are proportional to the number of subjects, with the smallest circle equal to one Y chromosome; connecting lines are proportional to the number of mutations.
doi:10.1371/journal.pone.0071390.g003

Table 2. Estimated Ages for Q Sub-Lineages in Native Americans.

Lineage (N[a])	Age ± SE (kya)
Q-total (431)	22.2±3.8
Q1a3a1-L54 (83)	24.4±3.4[b]
Q1a3a1-L54* (74)	23.6±3.8[c]
Q1a3a1a-M3 (348)	21.8±4.1
Q1a3a1a-M3* (334)	21.0±3.7[d]
Q1a3a1b-L191 (9)	3.6±1.6
Q1a3a1-DYS19-13.2 (6)[e,f]	9.4±5.8

SE, Standard Errors.
[a]excluding samples with duplicated loci and micro-variants (partial repeats).
[b]including PV4, 24.3±3.4 without PV4.
[c]including PV4, 23.7±3.7 without PV4.
[d]including or excluding PV2.
[e]one sample from this study and five from the YRHD database.
[f]DYS388 and DYS461 were not genotyped.
doi:10.1371/journal.pone.0071390.t002

chromosomes. The presence of Far East Asian samples in the central haplotype of the M3* network and of subjects from the Northwest territories of Canada in one- and two-step derivative haplotypes as well as the large haplotype sharing between Central and Southern America and the general north-to-south decreasing pattern of V and h, lead to the following considerations: i) haplogroup Q1a3a1a-M3 was probably the protagonist of a swift migration that originated a uniform pattern of distribution along the American continent; ii) the local differentiation identified in the Andean region due to long-term isolations of Q-M3 chromosomes supports an early arrival of this lineage in South America; iii) the northern territories of the Americas, despite subsequent arrival(s), retain the footprint of this first migration; and iv) even though M3 likely arose in Beringia, the extant North East Asian representatives could better represent the result of a back migration rather than be direct M3 Beringian descendants. With this regard, it is noteworthy that, after a Beringian population expansion concomitant with the entry to America, more recent (during the Holocene) contacts among Asian and American circum-arctic populations were proposed to explain the distribution of morphological characters [79]. The presence of M3 subjects in some Far East Asian groups could also be due to recent contacts (gene flows) with modern northern Native Americans. Their presence in populations such as coastal Chukchi [78], Evens [59] and Koryaks (present study), characterized by a subsistence economy which integrates land-based activities (reindeer herding or dog breeding) combined with marine animal hunting, gives support to this interpretation. Moreover, the presence in these North Eastern groups of chromosomes belonging to haplogroup Q1a-MEH2*, which characterizes the 4 ky-old Palaeo-Eskimo individual of Saqqaq culture (marine hunting), confirms the genetic continuity between modern and Palaeo-Eskimo populations. The estimated age of Q1a-MEH2* microsatellite variation obtained after having included our samples (Table 3) is in agreement with the scenario that the hypothesized crossing of the Bering Straits occurred 5.5 kya, independent from previous entry waves [39]. Accordingly, Q1a-MEH2, previously described in different indigenous groups from Far North-Eastern Russia [59,78,80] and Southeast Alaska [49], was not found in Central and South America.

Finally, our Mongolian samples (Table S5), which are from a geographic area lying over the Altai region of Mongolia (Western Mongolia) and the northern part of the country (including also the territory of Kazakh and Buryat ethnic groups, historically scattered between Balkhas and Baikal Lakes), share with southern Altaians and southern Siberian populations [51] the L54 sub-haplogroup Q1a3a1c-L330, and with northern Altaians the phylogenetically upstream Q-L53* observed in one subject (the only one from Eastern Mongolia) (Figure 3 and Table S5). These observations are in agreement with the hypothesis that the Native American ancestral population comes from somewhere in southern Siberia, between the Altai Mountains (which acted as a genetic and cultural wall) and the Lake Baikal region [16]. Thus, it is likely that the ancestral form of the L54 Y chromosomes was already comprised in the probable Asian source population(s), which, as supported by archaeological records [81], in prehistoric times and long before the peopling of the New Continent, moved eastwards during the Beringian standstill. In this setting, they evolved into the Native American Y-chromosome founding lineages (Q-L54* and Q-M3) [3], which began a rapid southwards diffusion following climatic ameliorations. In this context, it is noteworthy that the only Q sample from southeastern Europe was Ukrainian [82], which is now classified as M346*(xL53, L54, M3, L191, L330, L401, L323), suggesting that the spread of haplogroup Q westwards probably occurred via Central Asia prior to the occurrence of the L53 and L54 mutations.

In this complex scenario, our data (1) are in agreement with the southern Siberia origin of Native American ancestral populations and the differentiation of both Native American Q founding lineages in Beringia; and (2) support a concomitant arrival of the Q-M3 and Q-L54* founding lineages in Mesoamerica (Table 2 and 3) and, therefore, the model proposed by Sandoval et al. [65] where Mexico was the recipient of the first wave of migration. However, our data also (3) evidence different patterns of evolution for the two founding lineages in the Mexican plateau and the Andean area, which can be explained by local differentiations of the two founder lineages after their rapid arrival first in Mexico and then, along the Pacific coast, in the Andean region. The differentiations observed in Mexico and the Andean region can be mainly ascribable to demographic expansions triggered by the introduction of agriculture and the flourishing of the Great Empires.

Materials and Methods

The sample

We analyzed 463 Y chromosomes belonging to haplogroup Q mainly from the Sorenson Molecular Genealogy Foundation (SMGF – http://www.smgf.org) database (Table 1). Candidates for haplogroup Q Y chromosomes were identified and selected according to genealogical, geographic and STR haplotype information. As for Native Americans, in the initial SMGF sample of 15,000 DNA donors from 31 countries ranging from North to South America, we identified 4,037 potential Native American subjects with paternal ancestors' birthplaces in the Americas for at least three generations. Among these, 2,025 presented STR haplotypes associated with haplogroup Q, according to the prediction algorithms. As for Asians, the initial sample included 1,179 unrelated males: 1,145 subjects from different areas of Mongolia and 34 Koryaks from the Kamchatka Peninsula.

The haplogroup prediction process was carried out using the SMGF prediction algorithm drawn on data from a subset of samples genotyped for 37 STR *loci* and signature haplogroup markers. A second publicly available Y-chromosome haplogroup predictor algorithm [83] was used for validation purposes. Confidence scores were also evaluated.

Table 3. Estimated Ages for Q Sub-Lineages in Mexico, Andes, Mongolia and Far East Asia and, in Comparative Populations.

Lineage	Country/Area/Population Age ± SE (kya) (N^a)							
	Mexico	Andes[b]	Mongolia	Far East Asia[c,d]	Altaians[d,e]	Athapaskan(Na-Dené)[f,g]	Eskimo-Aleuts[f,g]	Indigenous communities-Andes[d,h]
Q-Total	21.1±4.4 (178)	23.4±3.9 (196)	14.5±2.9 (19)[i]	22.4±8.5 (20)	10.3±3.1 (45)			
Q1a-MEH2*				3.5±2.3 (9)				
Q1a3a-L53*					3.3±1.7 (25)			
Q1a3a1-L54* (xMexican cluster)	23.4±5.5 (49)[j] 19.3±3.3 (35)	17.7±2.5 (19)[k]				5.6±2.3 (5)[l]		
Q1a3a-M3* (xAndean cluster)	19.3±4.5 (121)	23.7±4.4 (161)[m] 20.5±4.3 (131)				3.4±1.6 (21)	7.4±3.4 (6)	
Q1a3a1b-L191	2.3±1.0 (8)							
Q1a3a1c-L330			6.5±2.5 (14)		2.9±1.2 (20)			
Q1a3a1a4-SA01		7.2±2.6 (5)[n]						5.9±2.1 (16)
Q1a3a1a5-M557		2.6±1.4 (7)						

[a] samples with duplicated alleles and partial repeats were excluded.
[b] Andes: Bolivia, Chile, Colombia, Ecuador, Peru.
[c] Koryaks (present study), Evens [59], Chukchi [78].
[d] DYS388 and DYS461 were not genotyped.
[e] Altai Republic [51].
[f] North America [49].
[g] DYS461 was not genotyped.
[h] Bolivia and Peru [56].
[i] Q-M120 (3) and Q-M25 (2) samples included, Q-L53* sample (1) not included because not genotyped for DYS461.
[j] 23.3±5.4 (48) without PV4.
[k] 20.7±3.7 (20) including outlier sample carrying DYS391 = 6 repeat.
[l] including one Mi'kmaq from Nova Scotia and one First Nation from British Columbia.
[m] with or without PV2.
[n] 15.4±6.5 (6) including outlier sample carrying DYS390 = 24 repeat; 18.5±7.7 (6) without DYS388 and DYS461; 9.1±3.0 (5) excluding outlier sample carrying DYS390 = 24 repeat and without DYS388 and DYS461.
doi:10.1371/journal.pone.0071390.t003

Ethics Statement

All experimental procedures and individual written informed consent, obtained from all donors, were reviewed and approved by the Western Institutional Review Board, Olympia, Washington (USA) and by the Ethics Committee for Clinical Experimentation of the University of Pavia, Board minutes of the 5th of October 2010.

STR and SNP genotyping

All Native American and Mongolian samples were typed for 37 Y-STRs (four of which, YCAIIa/b, DYS385a/b, DYS459, DYS464, are duplicated *loci*) according to the SMGF website.

All samples predicted to be K-M9 and Q-M3 were analyzed for the haplogroup Q signature marker M242 [84] and all the M242-positive samples were hierarchically analyzed for 22 downstream markers by RFLP and/or DHPLC and/or Sanger sequencing of the pertinent PCR fragment. Primers and methods used for each marker are provided in Table S1.

Haplogroup nomenclature is according to Dulik et al. [49,51] and updated in agreement with the YCC [85] rules (http://ycc.biosci.arizona.edu).

Statistical and phylogenetic analyses

Haplogroup frequencies were calculated by direct counting. Diversity parameters (number of shared haplotypes, haplotype diversity) were calculated only using the minimal Haplotype (minHt – DYS19, DYS389I, DYS389B, DYS390, DYS391, DYS392, DYS393, DYS385a/b – [86,87] disregarding the duplicated DYS385a/b *locus*. All these analyses were performed using the Arlequin v. 3.5 software [88]. Samples with atypical alleles (duplicated and micro-variant alleles) were not considered in the evaluation of diversity parameters.

Principal Coordinates Analysis (PCoA), on pairwise, individual-by-individual genetic distances, based on thirty-three *loci* Y-STR haplotypes, was performed using Excel software implemented by GenAlEx 6.4 software [89]. The YCAIIa/b, DYS385a/b, DYS459 and DYS464 *loci* were not included in the analyses because of the impossibility of assigning each allele to one or the other *locus* of these systems. One sample with duplicated DYS390 *locus* was also excluded from this analysis.

Phylogenetic relationships among haplogroup Q Y chromosomes were investigated through Median-Joining (MJ) networks [90] generated with Network 4.6.1 and Network Publisher 2.0 software (http://www.fluxus-engineering.com), after having processed data with the reduced-median method [91]. Networks were generated only using the minHt, by weighting each *locus* proportionally to the inverse of its repeat variance and disregarding the duplicated DYS385a/b *locus*.

The age of microsatellite variation (coalescent time or expansion time) within haplogroup Q and its sub-lineages was evaluated by the methodology of Zhivotovsky et al. [57] as modified according to Sengupta et al. [52]. In particular, it was estimated as the average squared difference in the number of repeats between all current chromosomes and the median haplotype, averaged over microsatellite loci and divided by the effective mutation rate of 6.9×10 per 25 years, with the SE computed over loci (details, in Appendix A of Sengupta et al. [52]). A microsatellite evolutionary effective mutation rate of 6.9×10^{-4} per generation (25 years) was used [52], since it is suitable for time depths, such as those explored in this study, where the elapsed time frame is ≥ 1000 years or ~ 40 generations [58].

Although caution should be used in the dating procedures because of the lack of a reliable mutation rate for Y-chromosome STR *loci* – also the use of pedigree-based mutation rate gave archaeologically inconsistent results [51] – estimated ages give us an idea concerning the temporal relationships existing among the different sub-lineages within the haplogroup.

Supporting Information

Table S1   Description of analyzed markers.
(XLSX)

Table S2   Thirty-three *loci* Y-STR haplotypes in Native American Q chromosomes.
(XLSX)

Table S3   Variance and haplotype diversity values for haplogroups L54 and M3.
(XLSX)

Table S4   Seven loci Y-STR haplotypes in the analyzed samples.
(XLSX)

Table S5   Mongolian sample description and haplogroup classification.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: VB OS. Performed the experiments: VB NM VG. Analyzed the data: VB VG OS AA UAP NA SRW NM AT. Contributed reagents/materials/analysis tools: OS AT AA. Wrote the paper: VB OS. Performed sample collection: JEG-P. Read and provided comments on the manuscript: VB VG UAP NA JEG-P SRW AA NM AT OS.

References

1. Galanter JM, Fernandez-Lopez JC, Gignoux CR, Barnholtz-Sloan J, Fernandez-Rozadilla C, et al. (2012) Development of a panel of genome-wide ancestry informative markers to study admixture throughout the Americas. PLoS Genet 8: e1002554.
2. O'Rourke DH, Raff JA (2010) The human genetic history of the Americas: the final frontier. Curr Biol 20: R202–207.
3. Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, et al. (2007) Beringian standstill and spread of Native American founders. PLoS One 2: e829.
4. Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, et al. (2008) The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PLoS One 3: e1764.
5. Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogo MR, et al. (2008) Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. Am J Hum Genet 82: 583–592.
6. Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, et al. (2009) Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Curr Biol 19: 1–8.
7. Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, et al. (2010) The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. Genome Res 20: 1174–1179.
8. Kumar S, Bellis C, Zlojutro M, Melton PE, Blangero J, et al. (2011) Large scale mitochondrial sequencing in Mexican Americans suggests a reappraisal of Native American origins. BMC Evol Biol 11: 293.
9. Bodner M, Perego UA, Huber G, Fendt L, Röck AW, et al. (2012) Rapid coastal spread of First Americans: novel insights from South America's Southern Cone mitochondrial genomes. Genome Res 22: 811–820.

10. de Saint Pierre M, Bravi CM, Motti JM, Fuku N, Tanaka M, et al. (2012) An alternative model for the early peopling of southern South America revealed by analyses of three mitochondrial DNA haplogroups. PLoS One 7: e43486.

11. de Saint Pierre M, Gandini F, Perego UA, Bodner M, Gómez-Carballa A, et al. (2012) Arrival of Paleo-Indians to the southern cone of South America: new clues from mitogenomes. PLoS One 7: e51311.

12. Perego UA, Lancioni H, Tribaldos M, Angerhofer N, Ekins JE, et al. (2012) Decrypting the mitochondrial gene pool of modern Panamanians. PLoS One 7: e38337.

13. Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF (2004) High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol Biol Evol 21: 164–175.

14. Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci U S A 93: 196–200.

15. Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipova L, et al. (1997) Y chromosome markers and Trans-Bering Strait dispersals. Am J Phys Anthropol 102: 301–314.

16. Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, et al. (1999) Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. Am J Hum Genet 64: 817–831.

17. Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, et al. (2002) High levels of Y-chromosome differentiation among Native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. Hum Biol 74: 761–789.

18. Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, et al. (1997) Y chromosome polymorphisms in native American and Siberian populations: identification of native American Y chromosome haplotypes. Hum Genet 100: 536–543.

19. Lell JT, Sukernik RI, Starikovskaya YB, Su B, Jin L, et al. (2002) The dual origin and Siberian affinities of Native American Y chromosomes. Am J Hum Genet 70: 192–206.

20. Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SP, et al. (2003) Y-chromosome evidence for differing ancient demographic histories in the Americas. Am J Hum Genet 73: 524–539.

21. de Acosta J, Mangan JE, Lopez-Morillas F, Mignolo WD (2002) Natural and moral history of the Indies. Durham, NC: Duke University Press.

22. Hrdlička A (1928) The origin and antiquity of the American Indian. Washington: U.S. Government Printing Office.

23. Martin PS (1973) The discovery of America: The First Americans may have swept the Western Hemisphere and decimated its fauna within 1000 years. Science 179: 969–974.

24. Haynes G (2002) The early settlement of North America: The Clovis Era: Cambridge University Press.

25. Goebel T, Waters MR, O'Rourke DH (2008) The late Pleistocene dispersal of modern humans in the Americas. Science 319: 1497–1502.

26. Fagan BM (2004) The great journey: The peopling of ancient America. Gainesville: University Press of Florida.

27. Greenberg (1986) The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. Curr Anthropol 27: 477–497.

28. Torroni A, Schurr TG, Yang CC, Szathmary EJ, Williams RC, et al. (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 130: 153–162.

29. Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, et al. (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53: 563–590.

30. Merriwether DA, Rothhammer F, Ferrell RE (1995) Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 98: 411–430.

31. Bonatto SL, Salzano FM (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc Natl Acad Sci U S A 94: 1866–1871.

32. Volodko NV, Starikovskaya EB, Mazunin IO, Eltsov NP, Naidenko PV, et al. (2008) Mitochondrial genome diversity in Arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am J Hum Genet 82: 1084–1100.

33. Bolnick DA, Bolnick DI, Smith DG (2006) Asymmetric male and female genetic histories among Native Americans from Eastern North America. Mol Biol Evol 23: 2161–2174.

34. Hunley K, Long JC (2005) Gene flow across linguistic boundaries in Native North American populations. Proc Natl Acad Sci U S A 102: 1312–1317.

35. Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, et al. (2012) Reconstructing Native American population history. Nature 488: 370–374.

36. Phillips C, Rodriguez A, Mosquera-Miguel A, Fondevila M, Porras-Hurtado L, et al. (2008) D9S1120, a simple STR with a common Native American-specific allele: forensic optimization, locus characterization and allele frequency studies. Forensic Sci Int Genet 3: 7–13.

37. Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, et al. (2009) Haplotypic background of a private allele at high frequency in the Americas. Mol Biol Evol 26: 995–1016.

38. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, et al. (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18: 830–838.

39. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463: 757–762.

40. Dillehay TD (1997) Monte Verde, a Late Pleistocene settlement in Chile: the archaeological context and interpretation. Washington, DC: Smithsonian Institution Press.

41. Hooshiar Kashani B, Perego UA, Olivieri A, Angerhofer N, Gandini F, et al. (2012) Mitochondrial haplogroup C4c: a rare lineage entering America through the ice-free corridor? Am J Phys Anthropol 147: 35–39.

42. Dillehay TD, Ramírez C, Pino M, Collins MB, Rossen J, et al. (2008) Monte Verde: seaweed, food, medicine, and the peopling of South America. Science 320: 784–786.

43. Jenkins DL, Davis LG, Stafford TW, Campos PF, Hockett B, et al. (2012) Clovis age western stemmed projectile points and human coprolites at the Paisley Caves. Science 337: 223–228.

44. Dillehay TD (1999) The late Pleistocene cultures of South America. Evolutionary Anthropology: Issues, News, and Reviews 7: 206–216.

45. Erlandson JM, Rick TC, Braje TJ, Casperson M, Culleton B, et al. (2011) Paleoindian seafaring, maritime technologies, and coastal foraging on California's Channel Islands. Science 331: 1181–1185.

46. Fix AG (2005) Rapid deployment of the five founding Amerind mtDNA haplogroups via coastal and riverine colonization. Am J Phys Anthropol 128: 430–436.

47. Dillehay TD (2009) Probing deeper into first American studies. Proc Natl Acad Sci U S A 106: 971–978.

48. Smith BD (1997) The initial domestication of Cucurbita pepo in the Americas 10,000 years ago. Science 276: 932–934.

49. Dulik MC, Owings AC, Gaieski JB, Vilar MG, Andre A, et al. (2012) Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. Proc Natl Acad Sci U S A 109: 8471–8476.

50. Wei W, Ayub Q, Chen Y, McCarthy S, Hou Y, et al. (2013) A calibrated human Y-chromosomal phylogeny based on resequencing. Genome Res 23: 388–395.

51. Dulik MC, Zhadanov SI, Osipova LP, Askapuli A, Gau L, et al. (2012) Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. Am J Hum Genet 90: 229–246.

52. Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, et al. (2006) Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am J Hum Genet 78: 202–221.

53. Al-Zahery N, Pala M, Battaglia V, Grugni V, Hamod MA, et al. (2011) In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. BMC Evol Biol 11: 288.

54. Grugni V, Battaglia V, Hooshiar Kashani B, Parolo S, Al-Zahery N, et al. (2012) Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. PLoS One 7: e41252.

55. Zhong H, Shi H, Qi XB, Duan ZY, Tan PP, et al. (2011) Extended Y chromosome investigation suggests postglacial migrations of modern humans into East Asia via the northern route. Mol Biol Evol 28: 717–727.

56. Jota MS, Lacerda DR, Sandoval JR, Vieira PP, Santos-Lopes SS, et al. (2011) A new subhaplogroup of native American Y-Chromosomes from the Andes. Am J Phys Anthropol 146: 553–559.

57. Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, et al. (2004) The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am J Hum Genet 74: 50–61.

58. Zhivotovsky LA, Underhill PA, Feldman MW (2006) Difference between evolutionarily effective and germ line mutation rate due to stochastically varying haplogroup size. Mol Biol Evol 23: 2268–2270.

59. Malyarchuk B, Derenko M, Denisova G, Maksimov A, Wozniak M, et al. (2011) Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. J Hum Genet 56: 583–588.

60. Neves WA, Araujo AG, Bernardo DV, Kipnis R, Feathers JK (2012) Rock art at the pleistocene/holocene boundary in Eastern South America. PLoS One 7: e32228.

61. Acuña-Alonzo V, Flores-Dorantes T, Kruit JK, Villarreal-Molina T, Arellano-Campos O, et al. (2010) A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. Hum Mol Genet 19: 2877–2885.

62. Tovo-Rodrigues L, Callegari-Jacques SM, Petzl-Erler ML, Tsuneto L, Salzano FM, et al. (2010) Dopamine receptor D4 allele distribution in Amerindians: a reflection of past behavior differences? Am J Phys Anthropol 143: 458–464.

63. Hünemeier T, Amorim CE, Azevedo S, Contini V, Acuña-Alonzo V, et al. (2012) Evolutionary responses to a constructed niche: ancient Mesoamericans as a model of gene-culture coevolution. PLoS One 7: e38862.

64. Gayà-Vidal M, Moral P, Saenz-Ruales N, Gerbault P, Tonasso L, et al. (2011) mtDNA and Y-chromosome diversity in Aymaras and Quechuas from Bolivia: different stories and special genetic traits of the Andean Altiplano populations. Am J Phys Anthropol 145: 215–230.

65. Sandoval K, Moreno-Estrada A, Mendizabal I, Underhill PA, Lopez-Valenzuela M, et al. (2012) Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. Am J Phys Anthropol 148: 395–405.

66. Bailliet G, Ramallo V, Muzzio M, García A, Santos MR, et al. (2009) Brief communication: Restricted geographic distribution for Y-Q* paragroup in South America. Am J Phys Anthropol 140: 578–582.

67. Martínez-González LJ, Saiz M, Alvarez-Cubero MJ, Gómez-Martín A, Alvarez JC, et al. (2012) Distribution of Y chromosomal STRs loci in Mayan and Mestizo populations from Guatemala. Forensic Sci Int Genet 6: 136–142.

68. Kayser M, Brauer S, Schädlich H, Prinz M, Batzer MA, et al. (2003) Y chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. Genome Res 13: 624–634.

69. Ascunce MS, González-Oliver A, Mulligan CJ (2008) Y-chromosome variability in four Native American populations from Panama. Hum Biol 80: 287–302.

70. Nuñez C, Baeta M, Sosa C, Casalod Y, Ge J, et al. (2010) Reconstructing the population history of Nicaragua by means of mtDNA, Y-chromosome STRs, and autosomal STR markers. Am J Phys Anthropol 143: 591–600.

71. Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, et al. (2004) Phylogeographic analysis of haplogroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. Am J Hum Genet 74: 1014–1022.

72. Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, et al. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. Am J Hum Genet 74: 1023–1034.

73. King RJ, Ozcan SS, Carter T, Kalfoğlu E, Atasoy S, et al. (2008) Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. Ann Hum Genet 72: 205–214.

74. Myres NM, Ritchie KH, Lin AA, Hughes RH, Woodward SR, et al. (2009) Y-chromosome short tandem repeat intermediate variant alleles DYS392.2, DYS449.2, and DYS385.2 delineate new phylogenetic substructure in human Y-chromosome haplogroup tree. Croat Med J 50: 239–249.

75. Baca M, Doan K, Sobczyk M, Stankovic A, Węgleński P (2012) Ancient DNA reveals kinship burial patterns of a pre-Columbian Andean community. BMC Genet 13: 30.

76. Ruiz-Linares A, Ortíz-Barrientos D, Figueroa M, Mesa N, Múnera JG, et al. (1999) Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. Proc Natl Acad Sci U S A 96: 6312–6317.

77. Toscanini U, Gusmão L, Berardi G, Gomes V, Amorim A, et al. (2011) Male lineages in South American Native groups: evidence of M19 traveling south. Am J Phys Anthropol 146: 188–196.

78. Regueiro M, Alvarez J, Rowold D, Herrera RJ (2013) On the origins, rapid expansion and genetic diversity of native Americans from hunting-gatherers to agriculturalists. Am J Phys Anthropol 150: 333–348.

79. González-José R, Bortolini MC, Santos FR, Bonatto SL (2008) The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. Am J Phys Anthropol 137: 175–187.

80. Bisso-Machado R, Jota MS, Ramallo V, Paixão-Côrtes VR, Lacerda DR, et al. (2011) Distribution of Y-chromosome Q lineages in Native Americans. Am J Hum Biol 23: 563–566.

81. Pitulko VV, Nikolsky PA, Girya EY, Basilyan AE, Tumskoy VE, et al. (2004) The Yana RHS site: humans in the Arctic before the last glacial maximum. Science 303: 52–56.

82. Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, et al. (2009) Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet 17: 820–830.

83. Athey TW (2006) Haplogroup prediction from Y-STR values using a bayesian-allele frequency approach. J Genet Geneal 2: 34–39.

84. Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, et al. (2004) Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 114: 127–148.

85. The Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12: 339–348.

86. Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, et al. (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110: 125–133, 141–129.

87. Pascali VL, Dobosz M, Brinkmann B (1999) Coordinating Y-chromosomal STR research for the Courts. Int J Legal Med 112: 1.

88. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564–567.

89. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28: 2537–2539.

90. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37–48.

91. Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141: 743–753.